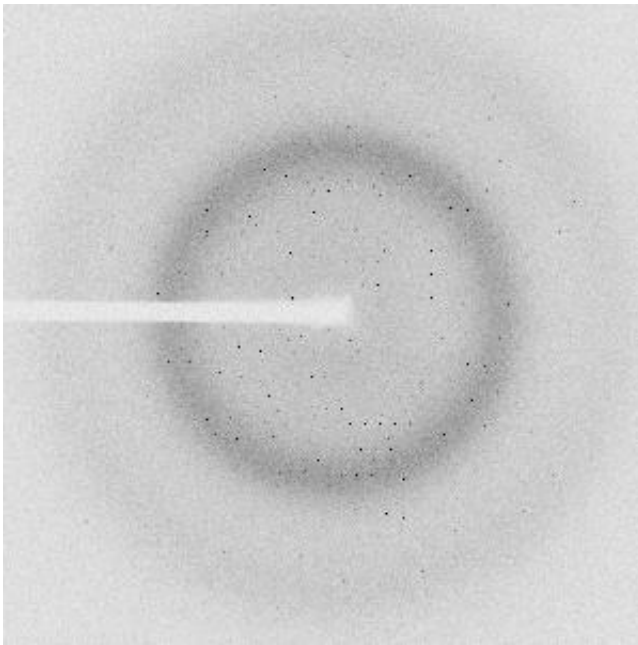


# A public database of macromolecular diffraction experiments

November 8 2016

---



Each dot forms from the constructive interference of X-rays passing through a crystal. The data can be used to examine the crystal's structure. Credit: M. Grabowski et al.

The reproducibility of published experimental results has recently attracted attention in many different scientific fields. The lack of availability of original primary scientific data represents a major factor contributing to reproducibility problems, however, the structural biology community has taken significant steps towards making experimental data available.

Macromolecular X-ray crystallography has led the way in requiring the public dissemination of atomic coordinates and a wealth of [experimental data](#) via the Protein Data Bank (PDB) and similar projects, making the field one of the most reproducible in the biological sciences.

The IUCr commissioned the Diffraction Data Deposition Working Group (DDDWG) in 2011 to examine the benefits and feasibility of archiving raw diffraction images in crystallography. The 2011-2014 DDDWG triennial report made several key recommendations regarding the preservation of raw [diffraction data](#). However, there remains no mandate for public disclosure of the original diffraction data.

The Integrated Resource for Reproducibility in Macromolecular Crystallography (IRRMCM) is part of the Big Data to Knowledge programme of the National Institutes of Health and has been developed to archive raw data from diffraction experiments and, equally importantly, to provide related metadata. The database [Grabowski et al. (2016). Acta Cryst. D72, 1181-1193, [DOI: 10.1107/S2059798316014716](#)], contains at the time of writing 3070 macromolecular diffraction experiments (5983 datasets) and their corresponding partially curated metadata, accounting for around 3% of all depositions in the Protein Data Bank. The resource is accessible at <http://www.proteindiffraction.org> and can be searched using various criteria via a simple, streamlined interface. All data are available for unrestricted access and download. The resource serves as a proof of concept and demonstrates the feasibility of archiving raw diffraction data and associated metadata from X-ray crystallographic studies of biological macromolecules.

Talking to a reporter about the project, team leader Wladek Minor said, "There is so much research underway that it can't all be published, and often the results of unsuccessful studies don't appear in the literature. I think the key to success is to know about unsuccessful experiments, we

want to know why they fail".

The goal of the project is to expand the IRRMC and include data sets that failed to yield X-ray structures. This could facilitate collaborative efforts to improve protein structure-determination methods and also ensure the availability of "orphan" data left behind by individual investigators and/or extinct structural genomics projects.

**More information:** Marek Grabowski et al, A public database of macromolecular diffraction experiments, *Acta Crystallographica Section D Structural Biology* (2016). [DOI: 10.1107/S2059798316014716](https://doi.org/10.1107/S2059798316014716)

Provided by International Union of Crystallography

Citation: A public database of macromolecular diffraction experiments (2016, November 8) retrieved 27 April 2024 from

<https://phys.org/news/2016-11-database-macromolecular-diffraction.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.